Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 00 (2019) 000–000

www.elsevier.com/locate/procedia

13th CIRP Conference on Intelligent Computation in Manufacturing Engineering, Gulf of Naples, Italy

Enabling reliable Visual Quality Control in Smart Factories through TSN

Jens Popper[a]*, Carsten Harms[b], Martin Ruskowski[b]

[a]*Technologie-Initative SmartFactory-KL, Trippstadter Str. 122, 67663 Kaiserslautern, Germany*
[b]*German Research Centre for Artificial Intelligence (DFKI), Trippstadter Str. 122, 67663 Kaiserslautern, Germany*

* Corresponding author. Tel.: +49-631-20575-3421; *E-mail address:* jens.popper@smartfactory.de

Abstract

Deep learning has been particularly successful in many fields such as computer vision in recent years. However, only few applications of Deep Learning can be found in the manufacturing context. Potentially overloading a computer network with the large amounts of data as well as limited computing power represent a big obstacle, especially for production sensitive data. To make Deep Learning applicable in production, these problems are described and a solution utilizing Time-Sensitive Networking Standards and transfer learning is developed. Then an exemplary application for the visual control of workpieces in ongoing production is implemented in a test factory.

Peer-review under responsibility of the scientific committee of the 13th CIRP Conference on Intelligent Computation in Manufacturing Engineering.

Keywords: Deep Learning, Artificial Intelligence, Visual Quality Control, Time-Sensitive Networking, Industry 4.0, Manufacturing, Smart Factory, Transfer Learning

1. Introduction

A globalized and connected world puts companies in an increasingly competitive environment. In addition, there is a change in purchasing behaviour. We are moving in more and more areas to a buyer-oriented market, in which customers demand products that are individually customized [1] to their needs and which are best delivered to them on the same day. This requires the integration of digital technologies into production. Such concepts are summarized in Germany under the term Industry 4.0 [2]. The use of cyber-physical production systems entails a networked manufacturing process in which large amounts of data are generated which are used for analysis purposes. An additional concept of Industry 4.0 is the use of modularized production systems with manufacturer-independent components.

The quality of the goods produced must also be guaranteed in the production of small batch sizes up to batch size 1. This poses special challenges for visual quality control, as the products can sometimes differ drastically from each other in their properties such as surface quality, shape, colour, integrated components and much more.

This paper shows a concept of how these visual quality control issues can be solved, and a system be integrated into a modular factory using new technologies and cross-manufacturer standards.

The structure is as follows: Chapter 2 names requirements that have been identified in such a context for visual quality control. Furthermore, Deep Learning concepts in the field of machine vision are briefly discussed. Some basics of the standards used for the network are also mentioned. In addition, the *SmartFactory*^{KL}, the test environment used for the implementation, is described. Chapter 3 deals with the detailed design of the individual concepts, technologies and components for the explicit use case and presents the results. Chapter 4 concludes with a discussion of the results.

2. Requirements and enabling technologies

Identifying objects with high variety is a big challenge for automated visual quality control methods [3]. Requirements that must be met by quality control in a production environment as described above are therefore as follows:

I. Quality control must be carried out continuously and reliably.
II. Quality control must be able to adapt flexibly to new product variants, which may also be produced simultaneously on the same production line.
III. Quality control must be scalable and applicable to various production processes.

These requirements also result from the design of the production facilities according to the lean principle, in which waste ("Muda") such as defective parts must be identified as quickly as possible and removed from the production process. However, over-engineering should still be avoided [4].

One way to meet these requirements for visual quality control is the use of artificial intelligence methods. These have led to several breakthroughs in recent years, especially in the field of computer vision [5]. In this approach we use Deep Learning, a sub-area of artificial intelligence [6].

However, the scalability requirement of the solution used must not only be considered from the point of view of the algorithms used. The infrastructure, and especially the network of the plant, must also be considered in production environments with large data streams. Notably, when using high quality video streams producing large amounts of data, one must ensure that the network of the plant cannot be overloaded. A naïve solution would utilize separate ethernet links for each application, however, such an approach requires retrofitting each time a new service is added. Alternatively, existing plants are retrofitted only once with Time-Sensitive Networking (TSN) equipment.

2.1. Short introduction to SmartFactory-KL

The *SmartFactory*^KL as first presented in [7] is a physical test-bed for novel production concepts, components and technologies for industrial applications and research projects. Based on the basic concepts of modularity, flexibility and manufacturer independence, this automated production line is divided into individual, mutually compatible production modules. These production modules are shown in Fig. 1.

Fig. 1 *SmartFactory*^KL Production modules

All production modules communicate with each other via standardized interfaces and communication protocols. High value is placed on manufacturer-independent standards such as OPC-UA. Thus, the process sequence can be changed quickly at any time or new processing steps can be integrated into the system and put into operation. This allows a high flexibility of the possible products to be manufactured.

The product to be manufactured is a business card holder. These can be individually assembled by the customers via an online product configurator and produced in the SmartFactory at the push of a button. An RFID chip is mounted in the bottom plate of the business card holder, on which all production information is stored (the product memory). The positions as well as completed machining processes are also stored, i.e. a digital twin is updated. RFID readers in each production module read out the product memory of the individual product to compare whether and which machining process must be carried out.

In addition, IP cameras were installed in each production module, which stream the machining process and the products within the modules.

2.2. Deep Learning for Computer Vision

The identification of surface defects and damage to a product plays an important role in the production environment. For a long time, this activity was mainly carried out manually by the company's employees, and only in the case of large batch sizes by automated systems. There it was mostly carried out by highly specialized software, which was adapted to the individual products with the help of expert knowledge. Especially for products with a high variety of variants, this form can only be implemented in a very complex way, which is why this implementation uses an Artificial Intelligence (AI) based form of Computer Vision for quality inspection.

Computer Vision refers to the use of computers and cameras to reproduce the human visual apparatus to identify objects, objects or other visual features. In recent years, the use of Convolutional Neural Networks (CNNs) has led to breakthroughs in the accuracy and flexibility of detection [8]. For example, by using CNNs, the accuracy of detection in the ImageNet challenge on the Top-5 could be increased from 73,8% to over 84,7% [9]. Current CNN based approaches achieve an accuracy up to 96,9% [10].

An advantage of Deep Learning-based approaches to the application of machine vision over traditional methods is that features do not need to be extracted manually [11]. These are therefore regarded as end-to-end models and do not require domain expert knowledge to be designed for a specific application.

These advantages make CNNs particularly interesting for visual quality control [17]. The network design used for the application presented in this paper is described in 3.1.

2.3. Short introduction to TSN

Time-Sensitive Networking (TSN) is a set of IEEE 802.1 published base standards and further projects being worked on by the TSN Task Group. Their main goal is to extend the nowadays in production extensively used Ethernet standard IEEE 802.3 with deterministic properties, such as upper latency bounds, low jitter (delay variance) and low packet loss.

These efforts meet many competing solutions (i.e. PROFINET, Ethernet/IP, EtherCAT), each realizing determinism over Ethernet differently with varying requirements. TSN unifies requirements to network equipment like switches and end points (i.e. PLCs, sensors, actors, cameras, etc.).

At the time of writing, the IEC/IEEE 60802 TSN Profile for Industrial Automation is not yet finished, thus we had to choose a suitable subset of standards considering both our use-case and supported TSN features by hardware manufacturers.

Switches must support IEEE 802.1Qbv (Time-aware scheduler) and IEEE 1588v2 (Precision Time Protocol, PTP) allowing to slice the available network bandwidth into timeslots to guarantee determinism to important traffic. End points must only support PTP to synchronize egress traffic with their respective timeslots. A software solution is precise enough for our use case.

3. Exemplary implementation into the SmartFactory

To meet the requirements listed in Chapter 2, the usage of Deep Learning technologies and a set of Ethernet standards is proposed. To test their suitability, they will be implemented exemplarily in the SmartFactory. For quality evaluation deep CNNs are used to continuously perform quality assurance (I) and to adapt it easily and flexibly to new products (II). Transfer learning [19] is also used for this purpose. TSN is used to ensure the reliability (I) and scalability (III) of the solution at the network level.

3.1. Description of the Convolutional Neural Network

The CNN used in this approach is based on the RetinaNet architecture presented in [12], consisting of ResNet-50 [13] Feature Pyramid Network shared box predictor with focal loss. The architecture of the network is shown in Fig. 2.

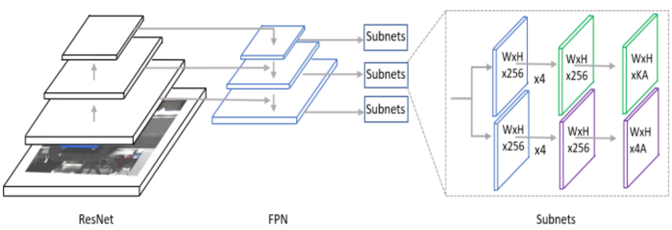

Fig. 2 One-Stage RetinaNet with ResNet, FPN, and subnetworks for anchor boxes [12]

This network was created using the TensorFlow Library [14] and pre-trained over 50,000 steps using the Common Objects in Context 14 (COCO14) dataset [15]. On this it achieved mean average precision (mAP) of 36.9. This trained network forms the basis for our implementation in the demo plant.

To adapt this network to detect the products in the production line, transfer learning is used [18]. This is applied by using the pre-trained network explained above and utilizing the low-level abstractions (such as edges) and training mostly the higher-level abstraction levels, such as the final object to be detected. This drastically reduces the amount of necessary training data which is required for the specific application and reduces the training time.

For this purpose, a further data set of 10 exemplary products was prepared, each consisting of 100 images. Furthermore, data augmentation in the form of cropping and horizontal and vertical flip was applied to a random number of images. In addition, a randomly generated noise effect was applied to a random selection of images. This data set was used to train the network of the COCO14 classification checkpoint to adapt it to the specific products.

Training was performed over 8000 steps on a NVIDIA GeForce GTX 1080Ti. From this created model the inference graph was exported and executed on a NVIDIA Jetson TX2. On this it could evaluate 9 frames per second (FPS).

3.2. Time-Sensitive Networking

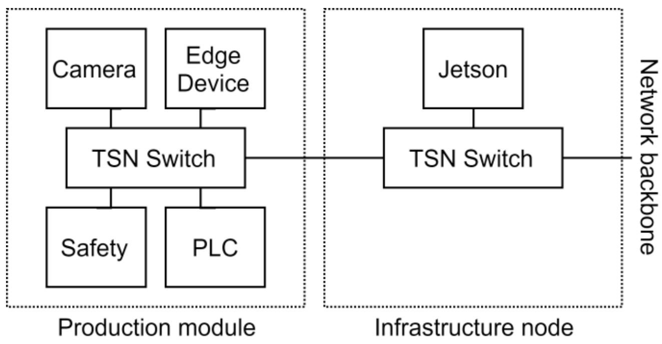

Fig. 3: Network topology utilizing a shared link

Shown in Fig. 3 is the relevant network topology. An edge device connected via a Hirschmann RSPE 35 TSN switch to the Cisco 6050 IP camera streaming at 30 FPS 1080p using a priority of 4. Utilizing the OpenCV library single frames of the camera stream are grabbed, zoomed into important image section and resized to the expected input size for the CNN. Afterwards, these single frames are sent to the NVIDIA Jetson with a higher priority of 5 of the single shared ethernet cable during the assigned timeslot described in Table 1 to fulfil the soft real-time requirements of our use case.

Table 1. Configured gate control list for IEEE802.1Qbv

Slot description	Priorities	Interval
Safety, Network Administration	6, 7	100 µs
Camera Streams	3, 4, 5	400 µs
Low Priority (Best Effort, etc.)	0, 1, 2	500 µs

Additionally, both hard real-time traffic (safety) and best effort traffic is utilizing the shared network link without any disturbances due to the time-aware schedulers of the Hirschmann switches.

We verified the correct operation and scalability by overloading the camera timeslot with an UDP flood generated by two laptops running the iperf3 tools over the shared link. As expected, the higher priority single frames within the same timeslot are not affected, because of strict priority-based scheduling. The same is true for the other two time slots due to time-aware scheduling. Thus, trying to add too many IP cameras to the network will only impair the video streams, but importantly not the prioritized single frames of each camera or any other network services in other time slots.

3.3. Implemented visual quality control

The inference graph of the Deep Learning network described in 3.1 was implemented in the *SmartFactory*^{KL} production line. The input for the evaluation of the computer vision algorithm is provided by a Cisco CIVS IPC 6050 IP camera installed on the top of a production module with view on the conveyor belt. The camera stream was then prioritized according to the scheduling specified in 3.2 and rescaled to a resolution of 640 x 640 pixels. The output of the evaluation represents the label about the recognized current processing state, which is compared in the system with the expected processing state for the respective product. In addition, a bounding box is created and placed over the camera stream to provide the operator with feedback on the identified products for traceability. The output generated in this way is shown in Fig. 4 as an example of a correctly and an incorrectly manufactured product.

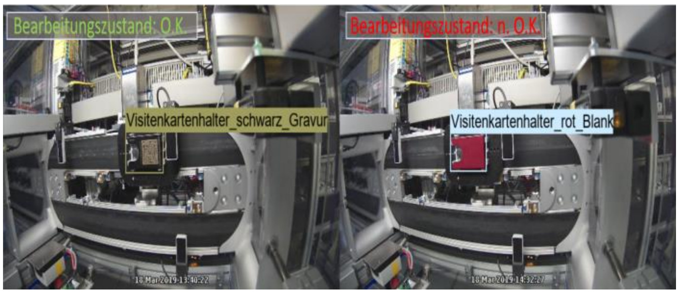

Fig. 4 Recognized process step of the product: correct (left) and incorrect (right)

4. Conclusion

In this paper a concept and a technical implementation of a Deep Learning based visual quality control is presented. By using pre-trained Deep Learning networks and transfer learning a high accuracy of the recognition is achieved and the number of required training images could be massively reduced, which enables a fast implementation and makes such an automatic recognition feasible even for very small lot sizes. The application and adaption to new products and new varieties is easy to do, which should enable even smaller companies without expert staff to implement and maintain these systems.

Using TSN switches and synchronizing an edge device within microseconds to the network time is sufficient to guarantee determinism required for reliable and continuous visual quality control. However, the manual configuration of the network is error prone and tedious.

We are aware of the future TSN "flow" concept [16] from talker to (multiple) listeners as well as an automatic schedule configuration via a central network controller (CNC); unfortunately, this functionality was not yet available for us. We strongly believe that these are necessary for practical implementation in future manufacturing plants.

Acknowledgements

This work has been partially funded by the European Commission through the FoF-RIA Project AUTOWARE: Autonomous, Reliable and Resilient Production Operation Architecture for Cognitive Manufacturing (No. 723909).

References

[1] Syam N.B, Ruan R, Hess J.D. Customized Products: A Competitive Analysis. In: Marketing Science, Vol. 24, No.4, 2005, p. 569-584.

[2] Kagermann H, Lukas W.-D, Wahlster W. Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. Industriellen Revolution. In: VDI-Nachrichten. April 2011.

[3] Bernal J, Vázquez D. Computer vision Trends and Challenges. 2013.

[4] Ohno T. Toyota Production System: Beyond Large-Scale Production. CRC Press, 1988.

[5] Hassaballah M, Hosny K. Recent Advances in Computer Vision: Theories and Applications. Springer International Publishing, 2019.

[6] Bengio Y, LeCun Y, Hinton G. Deep Learning. In: Nature. 512 (7553), 2015. p. 436-444.

[7] Zühlke D. SmartFactory – Towards a factory-of-things. In: Annual Reviews in Control 34. 2010. p. 129-138.

[8] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. In: Proceedings of the IEEE, Vol. 86 (11), 1998. p. 2278-2324.

[9] Krizhevsky A, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks. In: Advances in Neural Processing Systems 25, Vol. 1. 2012. p. 1097-1105.

[10] Imagenet lsvrc challenge, [Online]. Available http://image-net.org/challenges/LSVRC/. Last accessed 30.04.2019.

[11] Goodfellow I, Bengio Y, Courville A. Deep Learning. The MIT Press. 2017.

[12] Lin T.-Y, Goyal P, Girshick R, He K, Dollár P. Focal Loss for Dense Object Detection. Available Online: https://arxiv.org/abs/1708.02002. Last accessed: 30.04.2019.

[13] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. Available Online: https://arxiv.org/abs/1512.03385. Last accessed 30.04.2019.

[14] Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado G.S, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, Isard M, Jia Y, Jozefowicz R, Kaiser L, Kudlur M, Levenberg J, Mané D, Monga R, Moore S, Murray D, Olah C, Schuster M, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Viégas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y, Zheng X. Tensorflow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. Whitepaper, Available Online: http://download.tensorflow.org/paper/whitepaper2015.pdf. Last Accessed 30.04.2019.

[15] Lin T.-Y, Maire M, Belongie S, Bourdev L, Girshick R, Hays J, Perona P, Ramanan D, Zitnick C.L, Dollár P. Microsoft COCO: Common Objects in Context. Accessible Online: https://arxiv.org/abs/1405.0312 . Last Accessed: 30.04.2019.

[16] Whitepaper: Time Sensitive Networking: A Technical Introduction. Available Online: https://www.cisco.com/c/dam/en/us/solutions/collateral/industry-solutions/white-paper-c11-738950.pdf. Last Accessed: 30.04.2019.

[17] Popper J, Hermann J, Cui K, Bergweiler S, Weyer S, Ruskowski M, Wang M, Guang L, Hu Y. C, Kueh V, Wang D, Madigan M. C, Oppermann I, Ryu S.-M, Kim L, Wei S, Li R, Du X, Shang Y, Ding X,

Zhang G, Thonet G. Artificial intelligence across industries. IEC Whitepaper. 2018.

[18] Gao T, Stark M, Koller D. What Makes a Good Detector? – Structured Priors for Learning from Few Examples. In: European Conference on Computer Vision 2012. p. 354-367.

[19] Pan S. J, Yang Q. A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering Preprint. 2009. Available Online: https://www.cse.ust.hk/~qyang/Docs/2009/tkde_transfer_learning.pdf. Last accessed: 30.04.2019.